

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:04:29 ; Search time 15.0398 Seconds

(without alignments)
2160.495 Million cell updates/sec

Title: US-09-807-933B-1

Perfect score: 1836

Sequence: 1 MKFRTASSALLALALGTEM.....TYKEVTCPEKITAKTGCSSRK.338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	708.5	38.6	229	2 UC7308	cellulase (EC 3.2.
2	552.5	30.1	511	2 S10527	endoglucanase B pr
3	515.5	28.1	393	2 S59499	cellulase egli - s
4	241	13.1	471	1 A26160	cellulose 1,4-beta
5	241	13.1	471	1 A38979	cellulose 1,4-beta
6	188.5	10.3	410	1 S68153	cellulase (EC 3.2.
7	180	9.8	418	1 S28372	cellulase (EC 3.2.
8	173	9.4	438	1 S70602	cellulose 1,4-beta
9	153.5	8.4	1428	2 T09852	luciferin A - Califo
10	150	8.2	2704	2 S09118	G surface protein
11	149.5	8.1	475	2 S49886	probable membrane
12	147.5	8.0	2395	1 S50820	surface protein ty
13	147	8.0	242	2 S60143	cellulase (EC 3.2.
14	144	7.8	2718	2 A23475	G surface protein
15	143	7.8	1396	2 S36851	L-shaped tail fibre
16	142	7.7	860	2 S43846	xylinase B - rumen
17	141	7.7	962	2 S03818	carboxymethylcellu
18	138	7.5	495	2 B71360	hypothetical prote
19	137	7.5	1217	2 S52714	sericinB - silkmo
20	137	7.5	4776	2 B95206	cell wall surface
21	136.5	7.4	388	1 UC5461	cellulase (EC 3.2.
22	136	7.4	1275	2 T33369	hypothetical prote
23	135.5	7.4	881	2 S56032	probable membrane
24	135	7.4	302	2 S71334	acetyl xylan ester
25	135	7.4	388	2 S43920	cellulase (EC 3.2.
26	135	7.4	888	2 T46726	secreted acid phos
27	133	7.2	513	1 EUT01	cellulose 1,4-beta
28	133	7.2	513	2 S45380	cellulose 1,4-beta
29	132	7.2	286	2 C61615	sericin MG-2 - gre

30	132	7.2	513	2 S11439	cellulose 1,4-beta
31	130.5	7.1	1063	2 D86731	hypothetical prote
32	129	7.0	316	1 A38743	lacticin - human
33	128.5	7.0	1560	2 T02885	peroxisome prolif
34	127.5	6.9	537	2 B33485	spore coat protein
35	126.5	6.9	604	1 JP0001	glucan 1,4-alpha-g
36	126.5	6.9	856	2 T00349	Avicelase III - As
37	125.5	6.8	1217	2 T25894	hypothetical prote
38	125.5	6.8	2403	2 A53386	banro human
39	125	6.8	477	2 S53362	mucin 5AC (clone J
40	124.5	6.8	420	2 S53916	SUN4 protein precu
41	124.5	6.8	459	2 A25928	cellulase (EC 3.2.
42	124.5	6.8	786	2 T39585	ubiquitin protein
43	124	6.8	344	2 T40167	hypothetical prote
44	124	6.8	1131	2 T41144	hypothetical serin
45	123.5	6.7	402	2 E86185	hypothetical prote

ALIGNMENTS

```
RESULT 1
UC7308
cellulase (EC 3.2.1.4) - Scopulariopsis brevicaulis
N:Alternate names: endoglucanase I
C:Species: Scopulariopsis brevicaulis
C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: UC7308; PC7087
R:Nakatsuji, F.; Kawaguchi, T.; Takada, G.; Sumitani, J.; Moriyama, Y.; Arai, M.
Biosci. Biotechnol. Biochem. 64, 1238-1246, 2000
A>Title: Cloning and sequencing of an endoglucanase gene from Scopulariopsis brevicaulis
A:Reference number: UC7308
A:Accession: UC7308
A:Molecule type: DNA
A:Residues: 1-229 <NA>
A:Experimental source: strain TOF-1212
A:Accession: PC7087
A:Molecule type: protein
A:Residues: 21-37;149-164 <NA>
C:Genetics:
A:Gene: eg1
A:Introns: 147/3
C:Keywords: glycosidase; hydrolase

Query Match      38.6%; Score 708.5; DB 2; Length 229;
Best Local Similarity 57.5%; Pred. No. 9.8e-39;
Matches 122; Conservative 34; Mismatches 49; Indels 7; Gaps 5;

QY      129 ASGNQVTRPYMDCKKASCSMPGKANVS-SPVKSNKNDGVTALSDSN-AQSGCN-QGNSYM 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      21 ASGTGTTTRYPDCKKPCSCSPDKAPLSQGPMTCTDIND-NPLDDGLTSGCEPGGAYM 79

QY      186 CNDNOPWAVNDNLAVGFPAALAISSGGSRMCCSCELTFTSTSVAGKKVAVVNTNGDL 245
      : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
DB      80 CSHSPWAVDDELAVGMAAVNIGQGTESDMCCACYLEFTTGAVSGKKMIVQATNTGDL 139

QY      246 GSSTGHPFLQMPGCGVGFNCCSSQWGAHPNDKMSRYGSGISASDSSLPALQAGCM 305
      : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
DB      140 GNN---HPDIMPAGGVGIFNCGTQWCSPPMGWMDRQGVATRADCSFPALVAGCBW 196

QY      306 RFNMFNADNPMTYKEVTCPEKITAKTGCSSR 337
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      197 RFDWFGDNDPVSFRVECPALVOKSCQR 228

RESULT 2
S10527
endoglucanase B precursor - Pseudomonas fluorescens
C:Species: Pseudomonas fluorescens
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
C:Accession: S10527
R:Gilbert, H.J.; Hall, J.; Hazlewood, G.P.; Ferreira, L.M.A.
Mol. Microbiol. 4, 759-767, 1990
```

A;Title: The N-terminal region of an endoglucanase from *Pseudomonas fluorescens* subsp. *A*;Reference number: S10527; MUID:90355836; PMID:2117693
A;Accession: S10527
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-511 <GIL>
A;Cross-references: EMBL:X52615; NID:945497; PIDN:CAA3684.1; PID:945498
C;Superfamily: bacterial cellulose-binding domain homology; glycosidase GWGM domain hom
F;31-128/Domain: bacterial cellulose-binding domain homology <BCB>
F;180-217/Domain: glycosidase GWGM domain homology <GCG>
F;35-127/Disulfide bonds: #stacut predicted

Query Match 30.1%; Score 552.5; DB 2; Length 511;
Best Local Similarity 35.2%; Pred. No. 2.1e-28;
Matches 146; Conservative 46; Mismatches 128; Indels 93; Gaps 16;

QY 4 ITTASALLALALCTEWASAECKLYGQ-CGGKNNNGPFCES-----GSTCVSN 54
DB 98 IPIGSSVPEFVQ-CNNSSRAQVPAVTAICGGGSSAPSSVASSSSSVSTPSSS 156
QY 55 DYVSQCPLPGSSGKSSSAHKK-----TTAAHKTTT-----AAHKTTTAP 98
DB 157 SSVSSVPEGSSSSSVLTGCAQCMNYGTLPLCNNTSGMGYEDRSCVARTCSAQ 216
QY 99 AKTTTAAKASTPSSSS-----SSSGKYSVSGASG-----NGVTRYMD 140
DB 217 APGVIVSTSTPLSSSSSSSSSVASSSSSSSSSVSPIDGGCNGVATRYMD 276
QY 141 CCRASGSGPKANVS---SPYKCNKQGVTAISNSAQSGCNGNSIMCNDQPMVANDN 197
DB 277 CCKPHCGM--SANVPSLVSPLOSCANN-TFLSDVSVSGSDGGGYMCWDXLPFAVSPT 333
QY 198 LAYFAAALISGGESMCCSCFELTFT-----STVSAGKMWQVNTNGDGLG 246
DB 334 LAYVYATLS-SGD---VCGRCYLOFTGSSYNAPGPGSAAALAKMTIQAATIGIDV- 387
QY 247 SSTGAHFDLQMPGGVGI-FNGCSSQMGAPNDGWSRYGIGISA-----289
DB 388 --SGGQGDILVPGGGVAFNACSAQMGVSNALGAQVGFLLACKQQLGVNASLSQKSC 445
QY 290 --SQCSSL-----PSALQAGCKRPFNFKADNPSMTYKTECTKETATGCCR 337
DB 446 VLNRCDVFSRGLTLOQCGCTPAEFWFEADNPSLKYKEVCPAELTTSQGNR 500

RESULT 3

SS9499 cellulase eg11 - smut fungus (Ustilago maydis)
C;Species: Ustilago maydis (corn smut)
C;Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 28-May-1999
C;Accession: S59499
R;Schauwecker, F.; Manner, G.; Kahmann, R.
Biol. Chem. Hoppe-Seyler 376, 617-625, 1995
A;Title: Filament-specific expression of a cellulase gene in the dimorphic fungus *Ustilago*
A;Reference number: S59499; MUID:96145728; PMID:8590631
A;Accession: S59499
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-393 <SCCH>
A;Cross-references: GB:S81598; NID:g1478378; PIDN:AAB36147.1; PID:g1478379

Query Match 28.1%; Score 515.5; DB 2; Length 393;
Best Local Similarity 47.0%; Pred. No. 4e-26;
Matches 103; Conservative 31; Mismatches 64; Indels 21; Gaps 9;

QY 133 GVTTRYMDCKASGSGPKANVSFVSKNDGTAL---SDSNAGSCNGNSYMCNDN 189
DB 27 GMATRYMDCKLASWEGKAPVAPVADCKADGVTLLIDSKDPSGQSGCNGNFKMCSM 86
QY 190 QPW--AVVNDNLAYGPAALISGGESMCCSCFELTFT-----TSTSVAGKMWQVNTNGG 243
DB 87 QPFDDETPTLAFGF--GAFITGQESDTDCAFYAEFHDAGKAMKKNKLIPOVTINVG 144

QY 244 DLGSSYGAHFDLQMPGGVGI-FNGCSSQMGAPNDGWSRYGIGISSASDSSLPSALQAG 302
DB 145 DYVSQ---NFDPIPGGIGAFPKGCPAQMGVEASLMDQYGVXSATCSKLPRLQEG 201
QY 303 CKRRFN-WFKANDNPSM--TYKEVTCPEKITAKTGCSRK 338
DB 202 CKRRFSEW---GDNPLKGSPPKVPKCPKSLIDRSGCQRK 237

RESULT 4

A26160 cellulose 1,4-beta-cellulobiosidase (EC 3.2.1.91) II precursor - fungus (*Trichoderma reesei*)
N;Alternate names: exo-cellulobiohydrolase II
C;Species: *Trichoderma reesei*
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A26160; A26472
R;Chen, C.M.; Gritzali, M.; Stafford, D.W.
Bio/Technology 5, 274-278, 1987
A;Title: Nucleotide sequence and deduced primary structure of cellulobiohydrolase II from *Trichoderma reesei*

A;Reference number: A26160
A;Accession: A26160
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-471 <CHE>
R;Teeri, T.T.; Lehtovaara, P.; Kauppinen, S.; Salovuori, I.; Knowles, J.
Gene 51, 43-52, 1987
A;Title: Homologous domains in *Trichoderma reesei* cellulolytic enzymes: gene sequence and
A;Reference number: A26472; MUID:87248061; PMID:3596237

A;Accession: A26472
A;Molecule type: DNA
A;Residues: 1-471 <TEB>
A;Cross-references: GB:M16190; NID:g170540; PIDN:AAA34210.1; PID:g170541
C;Genetics:
A;Insertions: 31/2; 160/2; 243/1
C;Superfamily: cellulose 1,4-beta-cellulobiosidase II; fungal cellulose-binding domain hom
C;Keywords: glycosidase, hydrolase, polysaccharide degradation
F;31-62/Domain: fungal cellulose-binding domain homology <FCB>

Query Match 13.1%; Score 241; DB 1; Length 471;
Best Local Similarity 40.5%; Pred. No. 2e-08;
Matches 51; Conservative 21; Mismatches 46; Indels 8; Gaps 3;

QY 5 TTSSALLALALCTEWASAECKLYGQCGGKNNNGPFCESSTCKVSNDDYSCQLPBG 64
DB 8 TLATLATLAASVLEBRQA--CSSVWGCGGQWMSPTCCASGTCVSNDDYSCQLPGA 65
QY 65 SSGNKSSESAAHKTTTAAHKTTTAAHKTTTAAKSTPSSSSSGKYSY 124
DB 66 ASSSSSTRAA---STTSRVSPITSSSSATPPPGSTTRVPPVG--SGTATSGNPFVG 119
QY 125 VSGGAS 130
DB 120 VTPMAN 125

RESULT 5

A38979 cellulose 1,4-beta-cellulobiosidase (EC 3.2.1.91) II - fungus (*Trichoderma viride*)
C;Species: *Trichoderma viride*
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A38979
R;Wang, J.R.; Zhang, M.F.; Huang, T.
Acta Genet. Sin. 22, 74-80, 1995
A;Title: The primary structure of cellulobiohydrolase gene (CBH II) from *Trichoderma viride*
A;Reference number: A38979
A;Accession: A38979
A;Molecule type: DNA
A;Residues: 1-471 <WAN>
C;Genetics:
A;Gene: cbhII
A;Insertions: 31/2; 160/2; 243/1
C;Superfamily: cellulose 1,4-beta-cellulobiosidase II; fungal cellulose-binding domain hom
C;Keywords: glycosidase, hydrolase, polysaccharide degradation

F:31-62/Domain: fungal cellulose-binding domain homology <FCB>

Query Match 13.1%; Score 241; DB 1; Length 471;

Best Local Similarity 40.5%; Pred. No. 2e-08;

Matches 51; Conservative 21; Mismatches 46; Indels 8; Gaps 3;

QY 5 TTASSALLALGTEMASAECSKLYGQCGKNNNGPTCCESGTCVSNDDYSSQCLPSG 64
8 TATATLTLASVPLERQA--GSSWGGCGGQWMSGPTCCASGTCVSNDDYSSQCLPSG 65

QY 65 SGNKSESAHAKTTTAAHKTTTAAHKTTTAAKTTTAAKSTPNSSSSSSGKYS 124
Db 66 ASSSSSTRAA---STTSRVSPPTSSSSATPPPGSTTRKVPVG--SGTATYSGNPFG 119

QY 125 VSGGAS 130
Db 120 VTPMAN 125

RESULT 6

S68153

cellulase (EC 3.2.1.4) 3D precursor - Penicillium janthinellum

N/Alternate names: 3D endoglucanase 2; endo-1,4-beta-glucanase

C/Species: Penicillium janthinellum

C/Date: 11-Sep-1998 #sequence_revision 11-Sep-1998 #text_change 18-Jun-1999

C/Accession: S68153; S57950

R/Meritz, G.; Koch, A.; Henriessat, B.; Schulz, G.

Curr. Genet. 29, 490-495, 1996

A/Title: Endoglucanase II (EgII) of Penicillium janthinellum: cDNA sequence, heterologous

A/Reference number: S68153; PMID:96207475; PMID:8625430

A/Accession: S68153

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: mRNA

A/Residues: 1-410 <MER>

A/Cross-references: EMBL:X89564; NID:G984165; PIDN:CAA61740.1; PID:G984166

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995

C/Function:

A/Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as cel-

C/Pathway: cellulose degradation

C/Superfamily: Trichoderma cellulase III; fungal cellulose-binding domain homology

C/Keywords: glycosidase; hydrolase; polysaccharide degradation

F:1-21/Domain: signal sequence #status predicted <SIG>

F:2-410/Product: cellulase #status predicted <MAT>

F:26-57/Domain: fungal cellulose-binding domain homology <FCB>

Query Match 10.3%; Score 188.5; DB 1; Length 410;

Best Local Similarity 28.3%; Pred. No. 4.1e-05;

Matches 70; Conservative 31; Mismatches 77; Indels 69; Gaps 11;

QY 10 ALLAT---ALGTEMASAECSKLYGQCGKNNNGPTCCESGTCVSNDDYSSQCLPSG 66
11 ALLATGSLGQOTA-----WGQCGGCGGTGATTCVSGYCSQNNMTSQCCLGAA 62

QY 67 GKNKSESAHAKTTTAAHKTTTAAHKTTTAAKTTTAAKSTPNSSSSSSGK--YSA 124
Db 63 -----TSTSTTTTV-----YSTLKTTTSTSTSAAPGKXRFAG 98

QY 125 VSGGASGNGVTRRYWDCCASCSWPGKANVSPYKCNKDGVTALD--SNAQSCNGGNS 183
Db 99 VNIAGPFDGVT-----SGTODLSQVYDE--SYDGVNQNSHPVADT-----FNI 141

QY 184 YMCNDQNPVANDNLAYGFAPAAAIISGGGSRWCCSFELFTSTSVAGKKVVQVNV--- 240
Db 142 FRLPTGQVTFVNNL-----GSLDSNNFGKRYBOVGSGLSLGAYCIYDHNVAR 191

QY 241 -TGGDGL 246
Db 192 WNGGVIG 198

RESULT 7

S28372

cellulase (EC 3.2.1.4) III precursor - fungus (Trichoderma reesei)

N/Alternate names: endo-1,4-beta-glucanase; endoglucanase III

C/Species: Trichoderma reesei

C/Date: 11-Sep-1998 #sequence_revision 11-Sep-1998 #text_change 18-Jun-1999

C/Accession: S28372; S78516; S78517; S02626

R/Salohelmo, M.; Lehtovaara, P.; Penttilae, M.; Teeri, T.T.; Stahlberg, J.; Johansson, J.

Gene 61, 11-21, 1988

A/Title: EgIII, a new endoglucanase from Trichoderma reesei: the characterization of bo

A/Reference number: S28372; PMID:88255850; PMID:3384334

A/Accession: S28372

A/Molecule type: DNA

A/Residues: 1-418 <SALI>

A/Cross-references: EMBL:M19373; NID:G170548; PIDN:AAA34213.1; PID:G170549

A/Accession: S78516

A/Molecule type: mRNA

A/Residues: 1-418 <SALI>

A/Cross-references: GB:M19373; NID:G170548; PIDN:AAA34213.1; PID:G170549

A/Accession: S78517

A/Molecule type: protein

A/Residues: 22-23, 'X', 25-46, 'XX', 49-58 <SALI>

R/Stahlberg, J.; Johansson, G.; Petersson, G.

Eur. J. Biochem. 173, 179-183, 1988

A/Title: A binding-site-deficient, catalytically active, core protein of endoglucanase

A/Reference number: S02626; PMID:88185316; PMID:3356188

A/Accession: S02626

A/Molecule type: protein

A/Residues: 'XXX', 87-88, 'V', 90, 'X', 92-93, 'Y', 95-99 <STA>

A/Experimental source: culture filtrate

C/Genetics:

A/Gene: eg13

A/Introns: 110/2

C/Function:

A/Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as cr

C/Pathway: cellulose degradation

C/Superfamily: Trichoderma cellulase III; fungal cellulose-binding domain homology

C/Keywords: blocked amino end; glycoprotein; glycosidase; hydrolase; polysaccharide deg.

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-418/Product: cellulase #status experimental <MAT>

F:26-57/Domain: fungal cellulose-binding domain homology <FCB>

F:83-418/Product: cellulase core #status experimental <MAT2>

F:22/Modified site: pyrolydine carboxylic acid (Gln) (in mature form) #status experimen

F:124/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 180; DB 1; Length 418;

Best Local Similarity 23.5%; Pred. No. 0.00015;

Matches 68; Conservative 36; Mismatches 109; Indels 76; Gaps 10;

QY 8 SSALLALALGTEMASAECSKLYGQCGKNNNGPTCCESGTCVSNDDYSSQCLPSG 67
4 SVAPLLASILYGAIVAOQTWGGCGGIGMSGPTNCAPGACSTLNPYVACIPGAT 62

QY 68 NKSESAHAKTTTAAHKTTTAAHKTTTAAKTTTAAKSTPNSSSSSSG--KTSAYS 126
Db 63 -----ITT-----STRPGPTTTTATSTSTSTPTSSGVRPAGV 99

QY 127 GGASGNGVTRRYWDCCASCSWPGKAN-----VSPVSKCNKDGVTALSDSNAQS 176
Db 100 IAGPFGCTTD--GTCVTSKYVPLKNFTGSNNYPPGICQMOPVNEDEMT----- 149

QY 177 GCNGNSYMCNDQNPVANDNLAYGFAPAAAIISGGGSRWCCSFELFTSTSVAGKKVV 236
Db 150 -----FRLPVGMQVLYVNNL-----GGLNDSISIKYDVLVGCGLSGLAYCIV 192

QY 237 QVTN-----TGGDGLSS--TGAHFDLMPGGGVCIRFGSCSSQMGAPBDGW 279
Db 193 DIHNVARWNGGIITGGGPTNAQ-----TSLWSQLASKYASOSRNV 233

RESULT 8

S70602

cellulose 1,4-beta-cellubiosidase (EC 3.2.1.91) II precursor - cultivated mushroom

N/Alternate names: cellulase

C/Species: Agaricus bisporus (cultivated mushroom)

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

Db 1237 SS6SG6AS6SG6GAS6YDIUD6SG6DNKS6FI6DF6CI6F6LI6INC6C6ND6:6NC 1237

A;Gene: SGD:SIM1
A;Cross-references: SGD:S0001385; MIPS:YIL123w

A:Map position: 9L
 C:Superfamily: Saccharomyces NCA3 protein
 C:Keywords: transmembrane protein
 F:82-98/Domain: transmembrane #status predicted <tm>

Query Match 8.1%; Score 149.5; DB 2; Length 475;
 Best Local Similarity 20.4%; Pred. No. 0.015;
 Matches 71; Conservative 47; Mismatches 131; Indels 99; Gaps 13;

QY 5 TLASSLLALALCTEMASAECKLYGCGCKMNGPTCCSGSTCKVSNDRYSQCLPSPG 64
 DB 79 TSSAAGIAASIAVSSAALAKNEKISDAAA---SATASTSQASASSSSSSSATSTLES 134
 QY 65 SSGNKSSESAAHKT---TTAAHKTTTAAHKTTPAKKTTVAKASPNSSSSSSK 121
 DB 135 SVSSSEEAAPTSTVSTSTATOSASATKSTSTSTSTSTSTSTSTSTSTSTSTST 194
 QY 122 YSAVSGA-----SGNGVTRVY----- 139
 DB 195 SSSSSSGSISYGLADLDFSGPSEKFPQDTIPCDKFPGGQVISTDWIGEGMGVENTDTS 254
 QY 140 ----DC-----CKASCMPGKANVSFYKSNKDGVTALSDSNAGS---CNGNSYCNND 189
 DB 255 TGGSCKEGYSYSCQ--PGMSKTQWPSDQ-----PSDGRSVGLCKNGLYLSNTD 305
 QY 190 QPVAVDNLAAYGFAAALISGCGSRWCCSFELFTSTSVAG--KRVVGVNTNGDL--- 245
 DB 306 ADLCMGVBAAVVSKLSG-----VAICRTDYPGRNVIPRYEGSSSLP 354
 QY 246 -----GSSGAFHDLMPGGVGIFNGCSSQMGAPNDGMS 281
 DB 355 TVVDQDTYFTWEGKTSAGY--VYNNAGVSEVDCI--WGTSGSGIGN 398

RESULT 12

SS0820
 Surface protein type 51B - Paramedium tetraurelia
 C:Species: Paramedium tetraurelia
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 07-Dec-1999
 C:Accession: S50820
 R:Scott, J.; Leeck, C.; Roney, J.
 Nucleic Acids Res. 22, 5079-5084, 1994
 A:Title: Analysis of the micronuclear B type surface protein gene in Paramedium tetraurelia
 A:Reference number: S50820; MUID:95098630; PMID:7800503
 A:Accession: S50820
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-2395 <SCD>
 A:Cross-references: EMBL:U07603; NID:9467226; PID:AAA81947.1; PID:9467227
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1994
 C:Genetics:
 A:Genetic code: SGCS
 A:Introns: 472/3; 1310/3; 1821/3
 C:Superfamily: G surface protein

Query Match 8.0%; Score 147.5; DB 1; Length 2395;
 Best Local Similarity 22.1%; Pred. No. 0.091;
 Matches 83; Conservative 38; Mismatches 149; Indels 105; Gaps 18;

QY 22 SAACSKLYGCGCKMNGPTCC-----SGSTC-----KVSND-----YYSQC 60
 DB 1953 SKGACFR-YTSCSKLWNSDPTSCKMISNOCTGNSNCIGITLCSFNTDQGVSGYDAGC 2011
 QY 61 LPSGGSGNKSSESAAHKTTPA-----HKTTTAAHKTTTPAKKTTVAKASPNS 114
 DB 2012 IQGVPLNNSDPRVCPYISCAAFYTTTSDCOTASKKTTNG-----TTGCTALACGS 2066
 QY 115 SSSSSSKY-----SAVSGASGNGVTRVWDCCKASC-----SWPGKANVSPPVS 160
 DB 2067 YTTQACCYFNDKALILSGAITSIGTCT--WDTTASSCRQSCADLTGTHAACSSQLST 2124
 QY 161 CNKDGVTAL-----SDSNAGSGNG--GNSYMCNDNPAAVNDNLAYG--FAAAISGGG 211

DB 2125 CTSDETSCLVKACTSYTTQTACTAVGSDGICYMELASSTNNNTAKRCLLACADIONGT 2184
 QY 212 ESRWC-----CSCF--BLFTSTSVAGKMWVQVNTGDLGSSNG 250
 DB 2185 STNVCAVALISCVSNGTACIAKANCSTYTTKACSGGLDGI;CVFTQSTATATAAGTGTC 2244
 QY 251 AHFDLQMPGGVGIFNGCSS-----QMGAPNDGW--GSRVGGISASDCS-----SLP 296
 DB 2245 A-----LMTACTANSQVACQAAIDRCRSMWTAASGTGTTAVASKCATHTCATN 2292
 QY 297 SALQACCKRFRWFK 311
 DB 2293 QATNGACTRFLNWDK 2307

RESULT 13

S60143
 cellulase (EC 3.2.1.4) V precursor - fungus (Trichoderma reesei) (strain QM9414)
 N:Alternate names: endo-1,4-beta-glucanase V
 C:Species: Trichoderma reesei
 A:Variety: strain QM9414
 C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 17-Mar-1999
 C:Accession: S60143; S49043
 R:Salheimo, A.; Henriksas, B.; Hoffren, A.; Teleman, O.; Penttila, M.
 Submitted to the EMBL Data Library, May 1994
 A:Description: A novel small endoglucanase gene egls from Trichoderma reesei isolated by
 A:Reference number: S60143

A:Accession: S60143
 A:Molecule type: DNA
 A:Residues: 1-242 <SAL>
 A:Cross-references: EMBL:Z33381; NID:9485863; PID:9485864
 R:Salheimo, A.; Henriksas, B.; Hoffren, A.M.; Teleman, O.; Penttila, M.
 Mol. Microbiol. 13, 219-228, 1994
 A:Title: A novel, small endoglucanase gene, egls, from Trichoderma reesei isolated by e
 A:Reference number: S49043; MUID:95075308; PMID:7984103
 A:Accession: S49043
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 206-241 <SAM>
 A:Cross-references: EMBL:Z33381
 C:Genetics:
 A:Gene: egls
 A:Introns: 23/3; 45/3
 C:Function:
 A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce
 A:Pathway: cellulose degradation
 C:Superfamily: fungal cellulose-binding domain homology
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation
 F:1-17/Domain: signal sequence #status predicted <SPS>
 F:18-242/Product: cellulase #status predicted <MAT>
 F:210-241/Domain: fungal cellulose-binding domain homology <FCB>

Query Match 8.0%; Score 147; DB 2; Length 242;
 Best Local Similarity 67.6%; Pred. No. 0.011;
 Matches 23; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 29 LVGCGCKMNGPTCCSGSTCKVSNDRYSQCLP 62
 DB 209 LVGCGGAGWTPPTQAPCTCKVQNWYSQCLP 242

RESULT 14

A23475
 G surface protein - Paramedium primaurelia
 C:Species: Paramedium primaurelia
 C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 07-Dec-1999
 C:Accession: A23475
 R:Prat, A.; Katinka, M.; Caron, F.; Meyer, E.
 J. Mol. Biol. 189, 47-60, 1986
 A:Title: Nucleotide sequence of the Paramedium primaurelia G surface protein. A huge pro
 A:Reference number: A23475; MUID:87060934; PMID:3783679
 A:Accession: A23475
 A:Molecule type: DNA

